John von Neumann Institute for Computing

HEAVY: A High Resolution Numerical Experiment in Lagrangian Turbulence

Alessandra S. Lanotte, Federico Toschi

published in

Parallel Computing: Architectures, Algorithms and Applications ,
C. Bischof, M. Bücker, P. Gibbon, G.R. Joubert, T. Lippert, B. Mohr,
F. Peters (Eds.),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **38**, ISBN 978-3-9810843-4-4, pp. 721-728, 2007.
Reprinted in: Advances in Parallel Computing, Volume **15**,
ISSN 0927-5452, ISBN 978-1-58603-796-3 (IOS Press), 2008.

http://www.fz-juelich.de/nic-series/volume38

HEAVY: A High Resolution Numerical Experiment in Lagrangian Turbulence

Alessandra S. Lanotte[1] **and Federico Toschi**[2]

[1] Istituto per le Scienze dell'Atmosfera e del Clima, CNR, Str. Prov. Lecce-Monteroni
73100 Lecce, Italy and INFN, Sezione di Lecce, Italy
E-mail: a.lanotte@isac.cnr.it

[2] Istituto per le Applicazioni del Calcolo CNR, Viale del Policlinico 137, 00161 Roma, Italy and
INFN, Sezione di Ferrara, Via G. Saragat 1, I-44100 Ferrara, Italy
E-mail: toschi@iac.cnr.it

In the context of the DEISA Extreme Computing Initiative (DECI), we performed a state-of-the-art Direct Numerical Simulation (DNS) of heavy particles in an homogeneous and isotropic stationary forced three-dimensional turbulent flow at Reynolds number $Re_\lambda \simeq 400$. We report some details about the physical problem, computational code, data organization and preliminary results, putting them in the context of the present international research on Lagrangian Turbulence.

1 Introduction

Water droplets in clouds, air bubbles in water, pollutants or dust particles in the atmosphere, coffee stirred in a milk cup, diesel fuel spray into the combustion chamber are all examples of one single phenomenon: the transport of particles by incompressible turbulent flows. Particles can be neutrally buoyant, if their density matches the one of the carrying fluid, or they can have a mismatch in density. This causes light particles to be trapped in high vortical regions while heavy particles are ejected form vortex cores and concentrate in high strain regions. Recently, significant progress has been made in the limit of very heavy, pointwise particles[1], and some work has been done comparing experiments and numerical simulations[2]. Also, an International collaboration has been established at the purpose: the International Collaboration for Turbulence Research (ICTR)[3].

However, many questions are still unanswered: from very *simple* ones, e.g. concerning the correct equations for the motion of finite-size, finite density particles moving in a turbulent flow, to very complex ones, e.g. about the spatial correlations of particles with flow structures, or the inhomogeneous spatial distribution of inertial particles. This last is one of the most intriguing feature of these suspensions, i.e. the presence of particle clusters, in specific regions of the flow. It is easy to understand the relevance of such behaviour, since it can strongly modify the probability to find particles close to each other and thus influence their chemical, biological or mechanical interactions.

Here we present the results of high-resolution DNS of incompressible turbulent flows, seeded with millions of passive point-wise particles much heavier than the carrier fluid. The complete system -flow and particles-, can be characterized in terms of two dimensionless numbers, the Reynolds number Re and the Stokes number St. The former measures the turbulent status of the flow, while the latter measures the particle inertia. Our goal is to study aspects of the particle dynamics and statistics at varying both the flow turbulence, Re, and the particle inertia, St.

2 Physical Model and Computational Details

The numerical code, which solves the Navier-Stokes equations (2.1) for an incompressible three dimensional fluid, uses standard Pseudo-Spectral Methods[4]. The code relies on the use of Fast Fourier Transforms (FFT) for an efficient evaluation of the non-linear term of equations like (2.1) in Fourier space. We used the open source FFTW 2.1.5 libraries[5].
The Navier-Stokes equations describe the temporal evolution of a $3D$ viscous fluid, of viscosity ν, subject to the external force $\boldsymbol{f}$:

$$\frac{\partial \boldsymbol{u}(\boldsymbol{x},t)}{\partial t} + \boldsymbol{u}(\boldsymbol{x},t) \cdot \boldsymbol{\nabla} \boldsymbol{u}(\boldsymbol{x},t) = -\boldsymbol{\nabla} p(\boldsymbol{x},t) + \nu \Delta \boldsymbol{u}(\boldsymbol{x},t) + \boldsymbol{f}(\boldsymbol{x},t), \tag{2.1}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u}(\boldsymbol{x},t) = 0, \tag{2.2}$$

where the latter equation is the mass continuity for an incompressible flow. The force $\boldsymbol{f}(\boldsymbol{x},t)$, acting only at the largest scales of motion, keeps the system in a stationary state and is chosen to obtain a statistically homogeneous and isotropic flow.

The numerical code integrates the evolution for the vector potential field $\boldsymbol{b}(\boldsymbol{k},t)$, from which the velocity field can be obtained taking the curl ($\boldsymbol{u} = \boldsymbol{\nabla} \times \boldsymbol{b}$ or in Fourier space $\hat{\boldsymbol{u}} = \boldsymbol{k} \times \hat{\boldsymbol{b}}$). Thanks to this setting, the velocity field is automatically divergence free and in the evolution equation for $\boldsymbol{b}$ no pressure contribution is present. The code uses an explicit second order Adams-Bashforth (AB) scheme for the integration of the vector potential (the equation for $\hat{\boldsymbol{b}}$ can be derived from Eq.(2.1) with the viscous term exactly integrated).

The simplest form of the Newton equations for the heavy particle dissipative dynamics reads as follows:

$$\frac{d\boldsymbol{v}(\boldsymbol{X}(t),t)}{dt} = -\frac{1}{\tau_p}\left[\boldsymbol{v}(\boldsymbol{X}(t),t) - \boldsymbol{u}(\boldsymbol{X}(t),t)\right]. \tag{2.3}$$

These also are integrated with a second-order AB scheme.

The flow $\boldsymbol{u}(\boldsymbol{x},t)$ is integrated over a cubic domain, with periodic boundary conditions, discretized with a regular grid of spacing $\delta x = L/N$, where L is the size of the cubic box and N is the number of collocation points per spatial direction (see Table 1). Particle velocities $\boldsymbol{v}(\boldsymbol{X}(t),t)$ and positions $\boldsymbol{X}(t)$ are calculated by means of a trilinear interpolation.

3 Run Details and Performance

Within the DEISA project, we performed a DNS of Eqs.(2.1) at the resolution of $2048 \times 2048 \times 2048$ grid points, seeding the flow with millions of heavy particles. The flow and particle dynamics are fully parallel; a differential dumping frequency for particles is applied in order to have both highly resolved trajectories, and a sufficiently large statistical dataset (see details in the sequel). We used 512 processors, as the best compromise between having an acceptable number of processors (with respect to the global machine configuration), and memory limitations (our code requires about 500 Gbytes). Here we underline that, as the parallelization is performed in slices, we would have been anyhow limited to a number of processors less than or equal to 2048.
Regarding the monitoring of the numerical simulations, we divided the total run (about 400k cpu hours), in batch jobs ~ 5 hours long, for a total of about 155 jobs. At the end and beginning of each job a consistency check on the input/output files was performed. Indeed

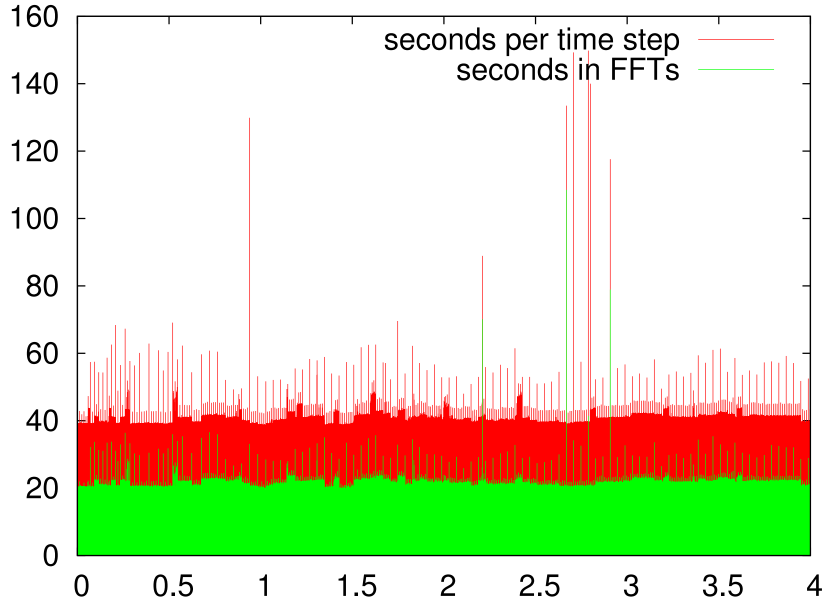

Figure 1. The computational performance of the DNS has been constantly monitored during the run (time of simulation in the horizontal axis is given in arbitrary units). The figure shows the total wall-clock time in seconds per time step, and the time spent (in seconds, per time step) for the FFTs. The relative weight of FFTs with respect to the whole loop in the time step is very close to 50%.

not only the program output was controlled for consistency, together with some overall physical indicators recorded very frequently, but also the sustained run performances were constantly monitored in order to understand the state of the machine during the runs. In Fig. 1 we show that performances were quite uniform during the run, with the presence of rare slowing down events which could take the wall-clock seconds per time step from the average 40 s to about 140 s i.e. a slow down of a factor $3 \div 4$. Obviously this phenomenon is related to hardware or networking issues.

An important observation which can be derived from Fig. 1 is the relative time spent in FFTs, with respect to the total wall-clock time step duration. As it can be seen, despite the relatively large load with $N_p = 2.1 \cdot 10^9$ particles (a load of 25%, i.e. one particle every four Eulerian grid points), almost 50% of the time is spent in Fourier transform and only the remaining 50% is spent in the temporal evolution of the Eulerian fields and of the Lagrangian particles. Probably this result can be understood as a consequence of the efficient and massively parallel way we use to integrate the particles: our conclusion is that for particle loads $\leq 25\%$ of the number of Eulerian collocation points, the computational cost for particle integration can be considered a fraction of the Eulerian cost.

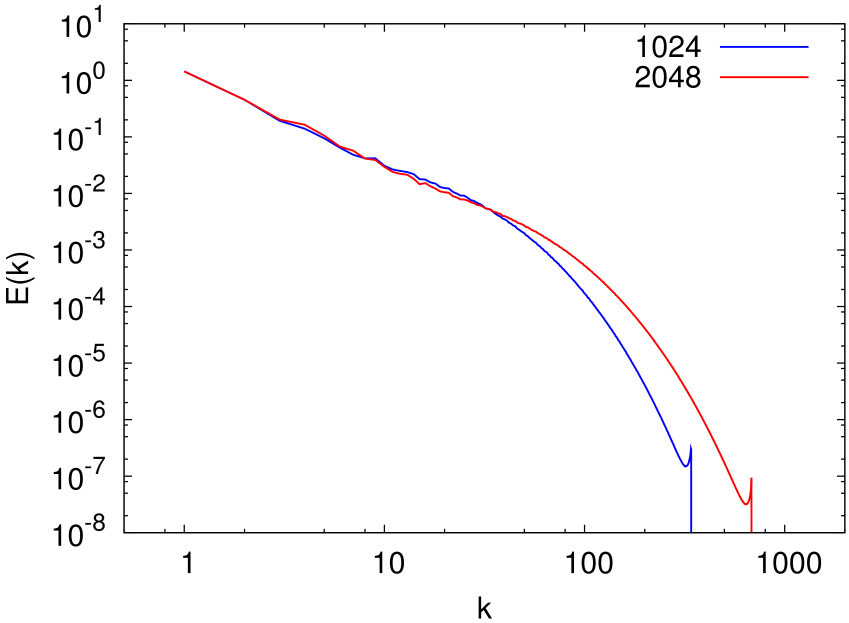

Figure 2. Log-log plot of the kinetic energy spectra $E(k)$ versus wavenumber k. The blue (lower) line is the spectrum of the starting configuration (coming from a thermalized simulation with $1024 \times 1024 \times 1024$ collocation points), while the red (upper) line is the thermalized spectrum for the current simulation with $2048 \times 2048 \times 2048$ collocation points. Other parameters, as the size of the system L and the force amplitude, are kept the same in the runs at the two resolutions.

3.1 Production Runs

Since the force injecting energy in the system acts only at the largest scales of motion, the time necessary for the turbulent flow to reach a statistically stationary state - with energy at all scales-, can be quite long. To reduce such thermalization time, the run was started from a stationary configuration obtained from a smaller resolution run with $1024 \times 1024 \times 1024$ grid points, performed during year 2004 at CINECA on 64 processors of the IBM SP4 machine[6,7]. The configuration at the lower resolution, expressed in Fourier space, was completed with zeroes at the largest wavenumbers and the run was started on the grid with $2048 \times 2048 \times 2048$ collocation points. Figure 2 shows the energy spectra, which measure the density of kinetic energy $E(k)$ per wavenumber $k = |\boldsymbol{k}|$, at the beginning and at the end of the thermalization stage. Once the Eulerian turbulent flow is stationary, we injected particles at initial random positions, with a uniform spatial distribution. We integrated the dynamics of a large number of particles with 21 different Stokes numbers. It would have been impossible to store the temporal evolution, with high temporal frequency-which means at about $\tau_\eta/10$-, for such a large number of particles. Hence we opted for a compromise. We dumped very seldom (every 4000 time steps, for a total of 13 dumps during the whole run duration), the information for all particles (see Table 1), together with the full Eulerian velocity field. This large amount of data can be very useful if one wants to study local particle concentrations -where large statistics is necessary-, or correlations

between particle properties/positions and Eulerian flow statistics.
Then, in order to investigate the temporal correlations along particle trajectories, we stored the particle information very frequently (typically every 10 time steps, i.e. roughly $\tau_\eta/10$). These data were stored only for a small subset of the total number of the particles (roughly $4 \cdot 10^6$ of the total number of particles, $N_p = 2.1 \cdot 10^9$). This mixed approach produced two distinct particle datasets, named **slow dumps** and **fast dumps**, and an Eulerian flow dataset, for a total disk occupation of 6.3 Tbytes.

At each particle dump, be it a slow or a fast one, we recorded all know information about the particle, i.e. name (a unique number useful in order to reconstruct particle trajectories), position, velocity, the fluid velocity at the particle position, the spatial gradients of the fluid at the particle position. These constitute a very rich amount of information about the statistical properties of heavy particles in turbulence. For some of the quantities we could measure, experimental Lagrangian measurements are extremely difficult because of limitations in the particles concentration, camera spatial resolution, etc.. In particular, to our knowledge, there are no measurements yet of the gradients of the fluid velocity at particle positions. So our large database will represent a very important test ground for theory and modelling.

As for particle inertia, we considered a wide range of values for the Stokes number St. Particles with very small inertia - e.g. microdroplets in clouds-, are very interesting to understand the statistics and dynamics of inertial particle in the limit of vanishing inertia, a limit which is thought to be singular. Particles with large Stokes number are also interesting since there are theoretical predictions[8], that still lack a direct assessment either by numerical or experimental measurements. Finally values around $St = 0.5$ are those for which preferential concentration is maximal[9], and this also needs further investigation. The physical parameters of our typical production runs are given in Table. 1.

Grid points	$N^3 = 2048^3$
Size of the system	$L = 2\pi$
Taylor Reynolds number	$Re_\lambda = 400$
Root-mean-square velocity	$u_{rms} = 1.4$
Kolmogorov time scale	$\tau_\eta = 0.02$
Fluid viscosity	$\nu = 4 \times 10^{-4}$
Time step	$dt = 1.1 \times 10^{-4}$
Kolmogorov scale	$\eta = 0.002$
Total number of particles	$N_p = 2.1 \times 10^9$
Stokes number	$St = 0.14, 0.54, 0.9, 1.81$
	$2.72, 4.54, 9.09, 18.18$
	$27.27, 36.36, 45.45, 63.63$

Table 1. Turbulent flow seeded with heavy particles: run parameters

4 Postprocessing and Pre-Analysis

4.1 Data Postprocessing

In order to attain the maximal performances during the numerical integration and I/O tasks, our program had the numerical integration of the particle dynamics fully parallelized among processors. This choice allowed us for a large particle load, but it implied that every processor had to dump its own particles at each time step. In order to reconstruct the trajectories of particles, from time to time, from one processor to another, we had to sort them during a postprocessing stage. We stored particles and Eulerian fields in the HDF5 dataformat[10]. For the Eulerian fields, the rationale of our choice was to store data in such a way that they could be easily analysed (for example in small cubes or in slabs) on small memory machines.

4.2 Pre-Analysis

We are currently running preliminary analysis, in particular we started by focusing on the Lagrangian Structure Functions, namely on moments of velocity increments at different time lags τ,

$$S_p(\tau) = \langle [v(t+\tau) - v(t)]^p \rangle = \langle (\delta_\tau v)^p \rangle. \tag{4.1}$$

In the previous expression, $v(t)$ is any of the three components of the particle velocity field, and the average is defined over the ensemble of particle trajectories evolving in the flow. As stationarity is assumed, moments of velocity increments only depend on the time lag τ. Morover, as our flow is statistically isotropic, different velocity components should be equivalent.
The presence of long spatial and temporal correlations, typical of any turbulent flow, suggests the existence of scaling laws in the inertial range of time $\tau_\eta \ll \tau \ll T_L$:

$$S_p(\tau) \sim \tau^{\xi(p)} \,. \tag{4.2}$$

In Fig. 3, we plot the second order Lagrangian structure function $S_2(\tau)$ for particles with three different values of inertia, $St = 0,\ 1.8,\ 63.6$. It is easy to see that while for the first two there are no detectable differences in the scaling behaviour, in the case of high inertia the shape of the curve dramatically changes. We remember that particles with very large inertia are very poorly sensible to the velocity fluctuations of the fluid velocity field, and their motion is almost ballistic. In other words, because the large value of their response time τ_p, for these particles fluid velocity fluctuations at time scales smaller than τ_p are filtered out: particle response time acts thus as a low pass filter for frequencies higher than $\sim 1/\tau_p$.

5 Concluding Remarks

In conclusion, within the DEISA Extreme Computing Initiative (DECI), we performed the most accurately resolved numerical simulation of Lagrangian turbulence worldwide. The

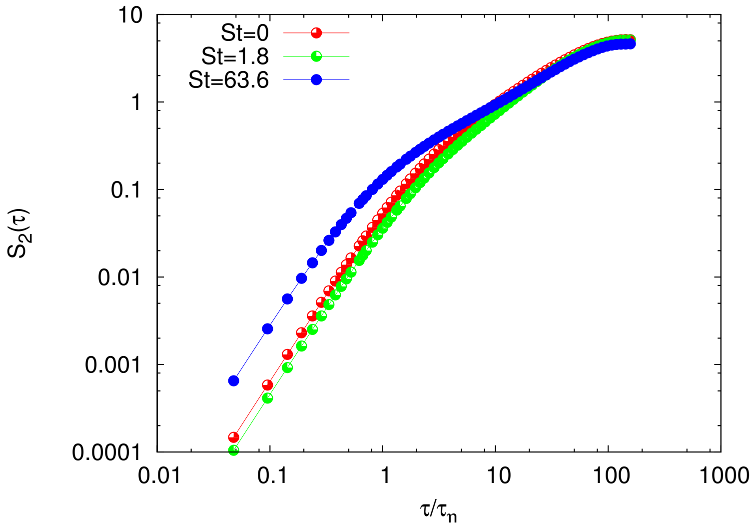

Figure 3. Log-log plot of the second order Lagrangian Structure Function $S_2(\tau)$ computed along the x component of the velocity field, versus time lags normalized with the Kolmogorov scale τ/τ_η. Here we have the curves for particles with three different values of the inertia: from below, the first curve is for tracer particles, $St = 0$; the central curve is for particles with $St = 1.8$, and the third curve is for particles with very large inertia, $St = 63.6$.

statistics of particles, and the space/time accuracy of the dynamics superseded those of all former studies.

Together with the handling of the output from this simulation, we faced the issue of establishing a standardized format for Lagrangian and Eulerian data. We relied upon HDF5 libraries and we implemented the dataformats in a way to ease as much as possible, from the coding and computational costs point of view, future data analysis.

The data from this study, after a preliminary period, will be made available to the whole scientific community at the International CFD database, iCFD database, accessible at the web site http://cfd.cineca.it and kindly hosted by CINECA[11].

Acknowledgements

We are indebted to Leibniz-Rechenzentrum for the extremely professional support, for the help provided during all the phases of the run: deployment, production and data postprocessing. In particular we would like to warmly thank Dr. Reinhold Bader.

We wish to thank Giovanni Erbacci, Chiara Marchetto and particularly Ivan Girotto from CINECA, for the continuos support and for the help during the delicate phase of data transfer and data storage. We thank Roberto Ammendola and Nazario Tantalo for the help with temporary data storage support at Tor Vergata University. We thank Luca Biferale, Massimo Cencini, Enrico Calzavarini and Jérémie Bec for constant support during the run preparation, testing and data analysis. We thank Claudio Gheller and Ugo Piomelli for many interesting discussions about data formats and, particularly, the HDF5 format. We thank CINECA for the hosting of the datasets produced from this run and former runs at the iCFDdatabase, http://cfd.cineca.it

Finally we thank the DEISA Consortium (co-funded by the EU, FP6 project 508830) for support within the DEISA Extreme Computing Initiative (www.deisa.org).

References

1. J. Bec, L. Biferale, M. Cencini, A. Lanotte, S. Musacchio and F. Toschi, *Phys. Rev. Lett.*, **98**, 084502, (2007).
2. L. Biferale, E. Bodenschatz, M. Cencini, A. S. Lanotte, N. T. Ouellette, F. Toschi and H. Xu. Lagrangian structure functions in turbulence: A quantitative comparison between experiment and direct numerical simulation (2007). http://it.arxiv.org/abs/0708.0311
3. `http://www.ictr.eu`
4. D. Gottlieb, S. A. Orszag, *Numerical analysis of spectral methods: Theory and applications* , SIAM (1977).
5. `http://www.fftw.org`
6. CINECA Keyproject for year 2004.
7. L. Biferale, G. Boffetta, A. Celani, B. Devenish, A. Lanotte and F. Toschi, *Physical Review Letters*, **93**, 064502, (2004).
8. J. Bec, M. Cencini and R. Hillerbrand, *Physica D* **226**, 11, (2007).
9. J. Bec, L. Biferale, G. Boffetta, A. Celani, M. Cencini, A. Lanotte, S. Musacchio, and F. Toschi, *J. Fluid Mech.*, **550**, 349, (2006).
10. `http://hdf.ncsa.uiuc.edu/products/hdf5`
11. `http://cfd.cineca.it`